Cross-Layer Geodiverse Protocol Stack for Resilient Multipath Transport and Routing using OpenFlow

Yufei Cheng*, Truc Anh N. Nguyen*, Md. Moshfequr Rahman*, Siddharth Gangadhar*,
and James P.G. Sterbenz*†‡
*Information and Telecommunication Technology Center
Department of Electrical Engineering and Computer Science
The University of Kansas, Lawrence, KS, 66045, USA
† School of Computing and Communications (SCC) and InfoLab21
Lancaster LA1 4WA, UK
‡Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
{yfcheng|annguyen|moshfequr|siddharth|jpgs}@ittc.ku.edu, jpgs@comp.{lancs.ac.uk,polyu.edu.hk}
www.ittc.ku.edu/resilinets

***Abstract*—This work describes a cross-layer resilient protocol stack for survivable network communications during regional challenges. The GeoDivRP routing protocol collects network statistics and calculates multiple geodiverse paths; it provides these geodiverse paths upstack to the resilient transport protocol, ResTP, for resilient multipath communications. ResTP provides multiple resilience modes to cope with different network failure conditions. A profile-based challenge model is used to better represent different challenge scenarios. Furthermore, the resilient protocol stack is implemented in network simulator *ns-3* and compared to Multipath TCP. Software-defined networking controller is proposed to implement the link failure detection module to increase the cross-layer protocol stack performance. By providing multiple d-distance separated paths, the protocol stack provides better path protection against regional challenges than MPTCP.**

***Index Terms*—path geodiversity; survivable routing heuristics; network resilience; Multipath TCP; OpenFlow**

I. Introduction and Motivation

The demands for Internet resilience have been increasing tremendously. Telecommunication networks are widely used for carrying Internet traffic and they rely heavily on physical infrastructure to maintain normal operation, and it is important to analyze their resilience to various faults and challenges [1]. Survivable optical networks under random edge and non-correlated failures have been a popular research domain [2], [3]. Recently, the research community has become more concerned about the potential damage caused by large-scale challenges and intentional attacks; efficient mechanisms have been proposed to mitigate their impacts [4], [5]. MPTCP (Multipath TCP) [6] is proposed to provide TCP with better network resource usage and more dependability against failures. However, none of this work has focused on a reliable cross-layer network architecture to cope with large-scale regional challenges.

Multipath routing has been adopted in the software-defined networking (SDN) environment [7]. MPTCP integration with the OpenFlow SDN has been evaluated [8], and several load balancing mechanisms have been proposed [9]. A datacenter network with a fat-tree topology has been used for OpenFlow research [10]. However, the research is mostly using datacenter networks and different challenge scenarios are not considered.

To deal with the aforementioned challenges, a novel flow-diverse Internet protocol stack has been proposed to provide network protection and resilience by taking advantage of multiple geodiverse paths. It provides network protection mechanisms by preallocating multiple geodiverse paths for each communicating node pair and employs erasure coding algorithms to take advantage of the multiple paths. The GeoDivRP routing protocol calculates the geodiverse path calculation by analyzing the network statistics collected from the link layer such as the failed nodes and links sets. The ResTP [11], [12] resilient transport protocol, establishes multiple flows between a pair of communicating hosts using the previously mentioned path set provided by GeoDivRP [13] for its data transmission. The dynamic interaction based on the current network condition between ResTP and GeoDivRP during a flow lifetime enhances the loss recovery process when the network suffers from a regional challenge or attack.

This paper compares the MPTCP [6] and the ResTP protocols. The simulation platform *ns-3* is used to compare the operation of both protocols in a geodiverse multipath challenge scenario. Section II describes the novel cross layer design within the flow-diverse resilient protocol stack. Section III details results comparing MPTCP to ResTP under challenges. Section IV concludes this paper. An extended version of this paper is published as a technical report [14].

II. Cross-layer design

ResTP [11] and GeoDivRP [13] form the transport and routing layer in the protocol stack. Knobs $\mathbb{K}$ are used by higher layers to influence the lower layer operation while dials $\mathbb{D}$ are the mechanisms for lower layers to provide instrumentation to the layers above. The application passes a service specification (ss) and threat model (tm) down to the transport layer protocol ResTP (resilient transport protocol). ResTP then requests GeoDivRP to calculate geodiverse paths that meet the requirement tuple $(k, d, [h, t])$, where k is the total number of geodiverse paths requested, d is the distance separation criteria, $[h, t]$ is the optional path stretch (number of additional hops for diverse paths) and skew (delay difference across paths) t constraint. Throughout this work, k is chosen as three for two main reasons. First, the node degree for the physical topologies used in this work is between three and four [15]. Second, a common spread used in erasure coding is three, which masks a single path failure. The resilient path is passed up to ResTP for resilient traffic communication.

ResTP establishes multiple transport flows for its data transmission by taking advantage of the geodiverse path set P_k and the traffic allocation information X_k provided by GeoDivRP; it chooses among its various reliability mechanisms to satisfy the specific application it is servicing according to the particular mission requirements. With the centralized view of the topology from the software-defined networking (SDN) environment, the failed node and link information can be passed upstack to notify GeoDivRP about the current network condition. With the updated network failure condition, GeoDivRP is be able to calculate multiple d-distance separated paths to provide improved resilience.

The OpenFlow SDN environment is employed to supplement the cross-layer design by providing the updated network condition. Our customized OpenFlow controller has the ability to detect any link failures and feed the information to GeoDivRP. Furthermore, since the OpenFlow controller has a centralized view of the network topology, it enhances the proposed protocol stack by providing real-time network statistics. Link failures are detected by the network monitor module, which is designed to be housed by the customized controller. The module provides network statistics, such as link failure information, and link congestion level, to GeoDivRP. GeoDivRP acts on this information and makes routing decisions such as the path to choose to get to the destination. Network statistics are acquired using OpenFlow discovery protocol (OFDP). The network devices advertise their link capacity and the controller constructs a centralized layer-2 network topology.

III. RESULTS

The *ns-3* [16] simulator is used to demonstrate the protocol stack's performance in face of regional challenges by comparing to MPTCP [17] using multiple node-disjoint paths. All the nodes in the topology are ResTP-GeoDivRP enabled, and path protection using multiple geodiverse paths is provided by GeoDivRP. Three geodiverse paths are used in the example topology for resilient routing using an erasure coding mechanism. We introduce the challenge profile study in this work and it provides a better understanding of how different regional challenges affect the network connectivity. As shown in Figure 1, the Sprint network [18] is studied with several challenge profiles. The movement for the Midwest profile is from the northwest to the southeast direction, while the hurricane profile moves from the northeast to the southwest.

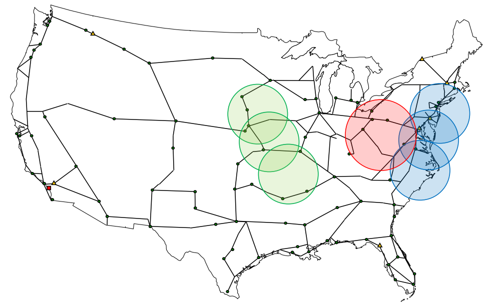


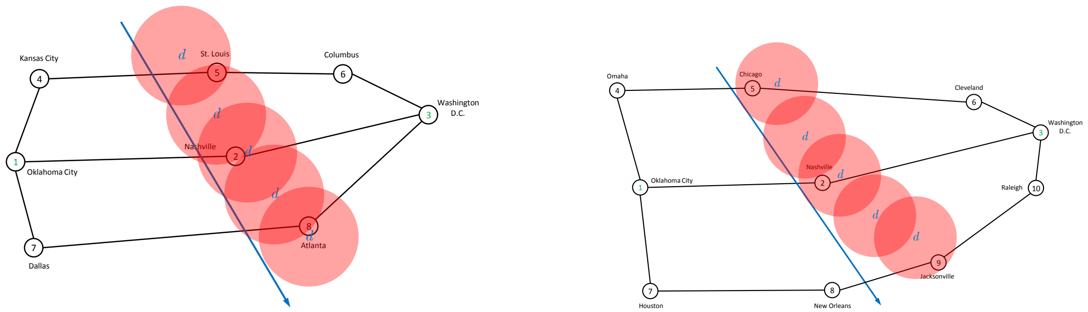
Fig. 1. Sprint network topology challenge profile

We use the Midwest profile in this work as it affects most of the shortest paths connecting the west and the east coast.

The Midwest challenge is applied in the Sprint network shown in Figure 2a and Figure 2b. Five challenge scenarios are included with failure radius d moving from the northwest to the southwest. The bandwidth on each link is 100 Mb/s, and higher data rates will be analyzed in future work. As observed from Figure 2a running MPTCP, the traffic originates from Oklahoma City to Washington D.C.; the paths for MPTCP is node-disjoints paths calculated using Suurballe's algorithm [19] and it cannot guarantee all the paths are geographically separated. In this experiment, MPTCP uses St. Louis, Kansas City, and Atlanta as its nexthops for the three paths. The challenge originates in St. Louis and moves towards Atlanta. The challenge radius is d, and since the paths are not d-distance separated, each challenge can affect two paths at the same time. When segments are divided onto three paths, the data packets lost due to path failure cannot be recovered.

On the other hand, GeoDivRP guarantees the paths are geographically separated, and therefore all the subflows created by ResTP are geodiverse; the protocol stack provides more throughput and resilience either using erasure coding or not. As shown in Figure 2b, GeoDivRP uses Omaha, Nashville, and Houston as its next hops. This guarantees that for any regional challenges with a radius no larger than the distance separation criteria d, at least 2/3 of paths survive; therefore, if the messages are properly coded, all the data packets can be delivered using the GeoDivRP and ResTP protocol stack. The traffic still originates from Oklahoma City to D.C.. The paths are provided by GeoDivRP using the iWPSP heuristic [13] with d-distance separation guaranteed. When the same Midwest challenge profile is applied to the network, each challenge can only affect 1/3 paths. When erasure coding is used on the data packets, the data can be delivered 100%.

Another challenge begins at 30 s in the northwest and moves to the southeast with each challenge lasting for 30 s. Figure 3 plots the average throughput in terms of Mb/s across the three paths against the simulation time. The throughput starts from zero and approaches 80 Mb/s at the beginning of the simulation until the first challenge occurs at 30 s. MPTCP does not guarantee geodiversity among the multiple paths; therefore, with the Midwest profile having a circular



(a) Sprint Midwest challenge using MPTCP

(b) Sprint Midwest challenge using ResTP

Fig. 2. Sprint network topology with multiple paths

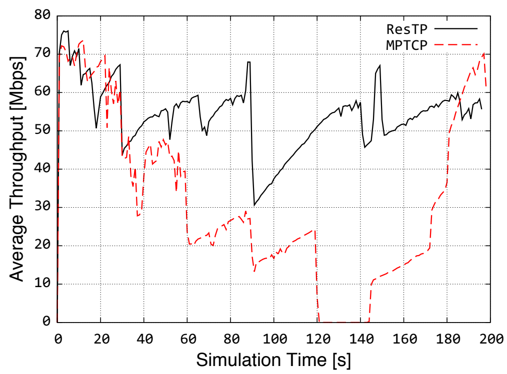


Fig. 3. ResTP throughput compared to MPTCP

radius d, each failure can take down two paths at the same time. With a real-time failure detection module planned to be part of our testbed, the protocol stack is predicted to respond to network condition challenges effectively and adapt accordingly with new geodiverse paths calculated. Overall, ResTP presents around 30% to 40% performance increase compared to MPTCP in face of regional challenges.

IV. Conclusion

This paper has presented the GeoDivRP-ResTP protocol stack and demonstrated its efficiency in bypassing the challenged regions and its improvement in terms of throughput compared to MPTCP. Through multiple geodiverse paths, the protocol stack meets the resilience requirement for different applications. With the real-time failure detection module planned to be part of our testbed, the protocol stack will take advantage of the provided link layer information and respond to network condition challenges, and adapt accordingly with new geodiverse paths calculated. Overall, the proposed protocol stack provides higher throughput and resilience than MPTCP in face of regional challenges.

Acknowledgment

The authors would like to thank the members of the ResiliNets group for discussions to led to this work. This research was supported in part by NSF grant CNS-1128122 and CNS-1219028.

References

[1] J. P. G. Sterbenz, D. Hutchison, E. K. Çetinkaya, A. Jabbar, J. P. Rohrer, M. Schöller, and P. Smith, "Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines," *Computer Networks*, vol. 54, no. 8, pp. 1245–1265, 2010.

[2] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, "Resilience of the Internet to Random Breakdowns," *Phys. Rev. Lett.*, vol. 85, pp. 4626–4628, November 2000.

[3] D. Magoni, "Tearing down the Internet," *IEEE J.Sel. A. Commun.*, vol. 21, pp. 949–960, September 2006.

[4] S. Neumayer and E. Modiano, "Network Reliability With Geographically Correlated Failures," in *IEEE INFOCOM*, pp. 1–9, March 2010.

[5] Y. Cheng, J. Li, and J. P. G. Sterbenz, "Path Geo-diversification: Design and Analysis," in *RNDM*, (Almaty), September 2013.

[6] A. Ford, C. Raiciu, M. Handley, and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple Addresses." RFC 6824 (Experimental), Jan. 2013.

[7] "OpenFlow: Multipath Proposal." http://archive.openflow.org/wk/index.php/Multipath_Proposal, 2011.

[8] R. van der Pol, S. Boele, F. Dijkstra, A. Barczyk, G. van Malenstein, J. Chen, and J. Mambretti, "Multipathing with mptcp and openflow," in *SCC*, pp. 1617–1624, Nov 2012.

[9] M. Bredel, Z. Bozakov, A. Barczyk, and H. Newman, "Flow-based load balancing in multipathed layer-2 networks using openflow and multipath-tcp," in *HotSDN*, HotSDN '14, (New York, NY, USA), pp. 213–214, ACM, 2014.

[10] T. N. Subedi, K. K. Nguyen, and M. Cheriet, "Openflow-based in-network layer-2 adaptive multipath aggregation in data centers," *Computer Communications*, vol. 61, pp. 58–69, 2015.

[11] T. A. N. Nguyen, J. P. Rohrer, and J. P. G. Sterbenz, "ResTP–A Transport Protocol for FI Resilience," in *Proceedings of the 10th International Conference on Future Internet Technologies (CFI)*, June 2015.

[12] J. P. Rohrer, A. Jabbar, and J. P. Sterbenz, "Path Diversification for Future Internet End-to-End Resilience and Survivability," *Springer Telecommunication Systems*, vol. 56, pp. 49–67, May 2014.

[13] Y. Cheng, M. T. Gardner, J. Li, R. May, D. Medhi, and J. P. Sterbenz, "Analysing GeoPath Diversity and Improving Routing Performance in Optical Networks," *Computer Networks*, vol. 82, pp. 50–67, May 2015.

[14] Y. Cheng, T. A. N. Nguyen, M. M. Rahman, S. Gangadhar, and J. P. Sterbenz, "Cross-Layer Geodiverse Protocol Stack for Resilient Multipath Transport and Routing using OpenFlow," ITTC Technical Report ITTC-FY2016-TR-66955-01, The University of Kansas, Lawrence, KS, January 2016.

[15] E. K. Çetinkaya, M. J. F. Alenazi, Y. Cheng, A. M. Peck, and J. P. G. Sterbenz, "On the Fitness of Geographic Graph Generators for Modelling Physical Level Topologies," in *RNDM*, (Almaty), pp. 38–45, September 2013.

[16] "The ns-3 Network Simulator." http://www.nsnam.org, July 2009.

[17] D. G. Fernández, "Multipath tcp in ns-3." https://github.com/dgomezunican/multipath-ns3.13, 2013.

[18] A. Cosner, J. P. Rohrer, M. J. F. Alenazi, and J. P. G. Sterbenz, "ResiliNets Topology Map Viewer Version 2." http://www.ittc.ku.edu/resilinets/maps/v2, January 2015.

[19] J. W. Suurballe, "Disjoint Paths in a Network," *Networks*, vol. 4, no. 2, pp. 125–145, 1974.